

s2/5/1

2/5/1

ALOG(R)File 351:Derwent WPI

) 2004 Thomson Derwent. All rts. reserv.

2388905 \*\*Image available\*\*

I Acc No: 1999-195012/199917

PX Acc No: N99-143282

Portable voice recognition device for telecommunication network -  
 generates feature vector data corresponding to voice of client from  
 microphone, which is then converted to acoustic signal for transmission  
 Patent Assignee: INT BUSINESS MACHINES CORP (IBMC ); IBM CORP (IBMC )  
 Inventor: KANEVSKY D; MAES S H; POON P S; PROCHILLO C  
 Number of Countries: 006 Number of Patents: 009  
 Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
11038992	A	19990212	JP 98152923	A	19980602	199917 B
5953700	A	19990914	US 97873079	A	19970611	199944
1229229	A	19990922	CN 98107970	A	19980508	200002
69315	A1	19991221	SG 981319	A	19980608	200006
99006431	A	19990125	KR 9816874	A	19980512	200014
99006628	A	19990125	KR 9820548	A	19980603	200014
394894	A	20000621	TW 98102629	A	19980224	200109
276846	B	20010115	KR 9816874	A	19980512	200206
6615171	B1	20030902	US 97873079	A	19970611	200359
			US 99776939	A	19990813	

Priority Applications (No Type Date): US 97873079 A 19970611; US 99776939 A  
 19990813

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
11038992	A	12	G10L-003/00		
5953700	A		G10L-005/06		
1229229	A		G10L-009/00		
69315	A1		H04M-011/06		
99006431	A		H04B-001/40		
99006628	A		H04L-012/28		
394894	A		G06F-003/18		
276846	B		H04B-001/40		Previous Publ. patent KR 99006431
6615171	B1		G10L-015/08		Cont of application US 97873079 Cont of patent US 5953700

Abstract (Basic): JP 11038992 A

NOVELTY - A microphone (100) converts sound containing client's voice and background noise, into an analog signal and is converted to digital signal. The digital signal is processed by a processor (120) to generate a feature vector data corresponding to voice of client. Vector data is converted into acoustic signal and is sent to destination through a sound coupler. DETAILED DESCRIPTION - An INDEPENDENT CLAIM is included for voice recognition procedure.

USE - For telecommunication network.

ADVANTAGE - Since the user compares the voice stored in server with the client's voice, confirmation of the client is effectively performed. DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of portable voice recognition device. (100) Microphone; (120) Digital signal processor.

Dwg.1/4

Keywords: PORTABLE; VOICE; RECOGNISE; DEVICE; TELECOMMUNICATION; NETWORK  
 GENERATE; FEATURE; VECTOR; DATA; CORRESPOND; VOICE; CLIENT; MICROPHONE;  
 CONVERT; ACOUSTIC; SIGNAL; TRANSMISSION

Derwent Class: P86; T05; W01; W04

International Patent Class (Main): G06F-003/18; G10L-003/00; G10L-005/06;

**THIS PAGE BLANK (USPTO)**

G10L-009/00; G10L-015/08; H04B-001/40; H04L-012/28; H04M-011/06  
International Patent Class (Additional): G10L-003/02; H04L-009/32;  
H04M-003/42  
Le Segment: EPI; EngPI

© 2004 Dialog Classic Web

**THIS PAGE BLANK (USPTO)**

JPA11-038992 which corresponds to USP 5,953,700

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-38992

(43) 公開日 平成11年(1999) 2月12日

(51) Int. Cl. <sup>6</sup>	識別記号	F I		
G10L 3/00	515	G10L 3/00	515	A
	531		531	L
	551		551	A
	561		561	H
3/02	301	3/02	301	A

審査請求 有 請求項の数26 OL (全12頁) 最終頁に続く

(21) 出願番号 特願平10-152923

(22) 出願日 平成10年(1998) 6月2日

(31) 優先権主張番号 08/873079

(32) 優先日 1997年6月11日

(33) 優先権主張国 米国 (US)

(71) 出願人 390009531

インターナショナル・ビジネス・マシー  
ズ・コーポレーション

INTERNATIONAL BUSIN  
ESS MACHINES CORPO  
RATION

アメリカ合衆国10504、ニューヨーク州  
アーモンク (番地なし)

(72) 発明者 ディミトリ・カネフスキー

アメリカ合衆国10562、ニューヨーク州オ  
シニング、スプリング・バレー・ロード  
1358

(74) 代理人 弁理士 坂口 博 (外1名)

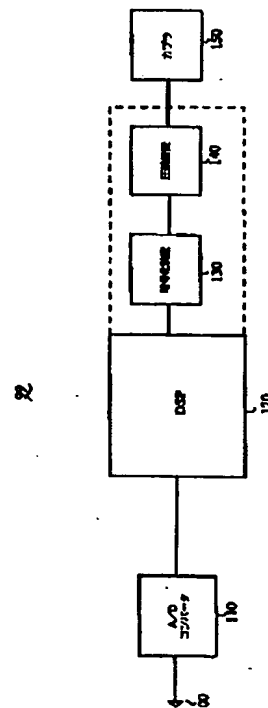
最終頁に続く

(54) 【発明の名称】 自動音声／話者認識サーバへのリモート・アクセス用ポータブル音響インタフェース装置、システム及び方法

(57) 【要約】

【課題】 話された音声及びバックグラウンド・ノイズを受信するマイクを持つポータブル音声信号前処理 (SSP) 装置、受信されたノイズを処理してフィーチャ・ベクトルを作成するデジタル信号プロセッサ (DSP)、及び通信チャンネルを通して転送するため通信装置に接続するカブラを提供する。

【解決手段】 自動音声／話者認識 (ASSR) サーバが通信チャンネルを通して、前処理された音声データを受信し、話された音声／話者を認識する。ポータブルSSP装置及びASSRサーバは、スマートカード、磁気カードまたは電子マネー・カードのPINコードの有効化、リセット、または変更をリモートに行うため使用できる。



## 【特許請求の範囲】

【請求項1】 音声、沈黙、バックグラウンド・ノイズの各信号を含むサウンドをアナログ信号に変換するマイクと、

アナログ信号をデジタル信号に変換するアナログ信号／デジタル・コンバータと、

前記デジタル信号から前記音声を表すフィーチャ・ベクトル・データを生成するデジタル信号プロセッサ(DSP)と、

前記フィーチャ・ベクトル・データを音響信号に変換し、遠隔地で前記音声を認識するために、前記音響信号を通信チャンネルを通して通信するために音響通信装置に接続する音響カプラと、  
を含む、ポータブル音声認識装置。

【請求項2】 前記フィーチャ・ベクトル・データを暗号化する暗号化装置を含む、請求項1記載の装置。

【請求項3】 前記フィーチャ・ベクトル・データを圧縮するデータ圧縮装置を含む、請求項1記載の装置。

【請求項4】 前記音響カプラは、前記遠隔地からの前記通信チャンネルを通した戻り信号を受信し、前記戻り信号を前記デジタル信号プロセッサによる処理のためにデジタル戻りデータに変換する手段を含む、請求項1記載の装置。

【請求項5】 前記デジタル信号プロセッサは、前記デジタル戻りデータを圧縮解除する手段を含む、請求項4記載の装置。

【請求項6】 前記デジタル信号プロセッサは、前記デジタル戻りデータを解読する手段を含む、請求項4記載の装置。

【請求項7】 前記通信チャンネルの伝達関数を予測する手段を含む、請求項1記載の装置。

【請求項8】 前記デジタル信号プロセッサは、前記遠隔地にて前記チャンネルに接続されたサーバからデータを受信し処理する手段を含む、請求項1記載の装置。

【請求項9】 前記デジタル信号プロセッサは、前記装置に固有のキー・データを含むデータを格納するメモリを含む、請求項1記載の装置。

【請求項10】 前記装置を囲む沈黙とバックグラウンド・ノイズに関連した通信チャンネル特性を把握する手段を含む、請求項1記載の装置。

【請求項11】 デジタル信号プロセッサ(DSP)に向かって話されたデジタル化された音声を処理し、前記音声を表すフィーチャ・ベクトル・データを生成する前記デジタル信号プロセッサと、

前記フィーチャ・ベクトル・データを表す信号を通信チャンネルを通して通信するために通信装置に接続するカプラと、

を含む、ポータブル音声信号プリプロセッサ(SSP)と、

前記通信チャンネルを介して前記ポータブル音声信号プリ

プロセッサから転送された前記フィーチャ・ベクトル・データを表す前記信号を受信し、前記音声または話者を認識するために前記信号を処理するために、前記通信チャンネルとリモートに接続された自動音声／話者認識(ASSR)サーバと、

を含む、音声認識システム。

【請求項12】 前記自動音声／話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を処理し、処理された信号を認証データの前記格納モデルと比較することによって有効な加入者を認証する認証データ格納モデルを含む、請求項11記載のシステム。

【請求項13】 前記自動音声／話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を処理し、処理された信号をデータベースと比較することによって音声を認識するためのボキャブラリの前記データベースを含む、請求項11記載のシステム。

【請求項14】 前記自動音声／話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を処理することによってユーザのパスワードを認証する、請求項11記載のシステム。

【請求項15】 前記自動音声／話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を処理することによって呼び出し側を識別し、前記信号は、前記呼び出し側によって前記ポータブル音声信号プリプロセッサに話された音声から発する、請求項11記載のシステム。

【請求項16】 前記ポータブル音声信号プリプロセッサは、前記自動音声／話者認識サーバから戻り信号を受信し処理する手段を含む、請求項11記載のシステム。

【請求項17】 前記戻り信号は変更された認証データ及びプロンプトの1つを含む、請求項16記載のシステム。

【請求項18】 前記ポータブル音声信号プリプロセッサは、前記自動音声／話者認識サーバから受信されたデータを含むデータを格納する関連メモリを含む、請求項11記載のシステム。

【請求項19】 前記通信チャンネルを通した前記ポータブル音声信号プリプロセッサと前記自動音声／話者認識サーバの間の接続の伝達関数を予測する手段を含む、請求項11記載のシステム。

【請求項20】 前記カプラは、前記フィーチャ・ベクトル・データを音響信号に変換する音響カプラであり、前記通信チャンネルは音響通信チャンネルである、請求項11記載のシステム。

【請求項21】 前記ポータブル音声信号プリプロセッサと前記自動音声／話者認識サーバは、通信セッション全体で話者認識を定期的に実行する手段を含む、請求項11記載のシステム。

【請求項22】 クライアント／サーバ・システムの有効

なクライアントを音声により認証する方法であって、登録データの複数のモデルを自動音声認識 (ASSR) サーバのメモリに格納するステップと、通信チャネルを通して前記自動音声認識サーバから遠く離れて配置されたポータブル音声信号前処理 (SSP) 装置から前記自動音声認識サーバに接続するステップと、前記登録データを前記ポータブル音声信号前処理装置に話すステップと、前記ポータブル音声信号前処理装置の前記登録データを前処理してフィーチャ・ベクトルを作成するステップと、前記フィーチャ・ベクトルを表す信号を前記通信チャネルを通して前記自動音声認識サーバにより受信するステップと、前記自動音声認識サーバから受信された前記信号を処理し、処理された信号を前記有効なクライアントを認証するために前記格納モデルと比較するステップと、を含む、方法。

【請求項 2 3】変更されたパスワード・データを前記自動音声認識サーバから前記ポータブル音声信号前処理装置に転送することによってユーザ・パスワードを変更するステップを含む、請求項 2 2 記載の方法。

【請求項 2 4】変更されたスマートカード・データを前記自動音声認識サーバから前記ポータブル音声信号前処理装置に転送することによってスマートカードの PIN を変更するステップを含む、請求項 2 2 記載の方法。

【請求項 2 5】前記変更されたスマートカード・データは前記自動音声認識サーバによって暗号化され、前記ポータブル音声信号前処理装置は前記スマートカードの PIN を変更するために前記変更されたスマートカード・データを解読する、請求項 2 4 記載の方法。

【請求項 2 6】変更された PIN データを前記自動音声認識サーバから磁気カード・ライタと前記ポータブル音声信号前処理装置に転送することによって磁気カードの PIN を変更するステップを含む、請求項 2 2 記載の方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】本発明は、自動音声／話者認識 (ASSR, Automatic Speech/Speaker Recognition) に関し、特に電話回線等の通信チャネルを通して ASSR サーバにリモートにアクセスするためのポータブル音声信号前処理装置またはインタフェースを用いた ASSR 法に関する。

【0 0 0 2】

【従来の技術】中央のサーバまたはサービス・プロバイダのリソースを多数のクライアントまたは加入者がアクセスする代表的なクライアント／サーバまたは加入者／サービス・プロバイダのシステムでは、何らかの形のク

ライアント／加入者認証法により、クライアント／加入者が有効なシステム・ユーザかどうか確認される。このようなシステムの多くはコードをキー入力するか、またはテキストをオペレータに通信することによって PIN、キーワードまたはパスワード等の認証コードをやりとりする。いくつかのシステムの認証コードは音声による。つまりコードはサーバに対して声で入力される。サーバ (この場合は ASSR サーバ) はその発音を認識し、記憶された有効な音声認証コードと比較することで、そのユーザが有効なクライアントであることを確認する。

【0 0 0 3】ASSR サーバ／クライアント・システムでは、クライアントが電話回線等の通信チャネルを通して音声により認証コードをリモートに通信すれば都合がよい。つまり電話回線が利用できるならどのクライアントもサーバにアクセスできるようにすることである。電話回線は地上回線またはセルラ回線等である。セルラの場合は、サーバへのアクセスは完全にポータブル、つまりいつでもどこでもセルラ・フォンとセルラ接続が利用できる。

【0 0 0 4】上に述べたような電話 ASSR システムの運営上の困難には、1) 電話回線で送られる音声データの劣化により正確さに欠ける、及び 2) 呼び出し側が公衆電話からかけているとき、車を運転しているとき等、ユーザがかけている電話の位置に応じて、ユーザ側でバックグラウンド・ノイズ特性が変化するということが含まれる。いずれの状況でもデータまたは信号の整合性が失われ、よって音声／話者の認識精度が大きく低下する。

【0 0 0 5】データ及び認識の精度が失われるこの問題は、クライアント側で信号が電話回線を通してサーバに送られる前に、音声信号前処理 (SSP, Speech Signal Preprocessing) が実行される場合は減らすことが、またはなくすることができる。SSP は転送デバイス、環境、話者及び通信チャネルの音響特性の把握を含む。SSP 情報は ASSR サーバによって処理され、基準が設定され、対応するデコード・モデルとアルゴリズムが選択され、ワード・エラー・レート小さくするため、または話者認識を正確に行うためにチャネル伝達関数とバックグラウンド・ノイズをモデリングすることによって、話者が認識されるか、または音声デコードされる。しかしユーザ側で SSP を実行するには、SSP ソフトウェアを持つコンピュータを含めた SSP 機器が必要になる。このような SSP 機能は、一般的には現在の標準的な電話や NC (ネットワーク・コンピュータ) にはない。

【0 0 0 6】従って、軽量コンパクトで持ち運びに便利であり、任意の電話またはデータ通信デバイスに接続でき、通信チャネルを通して ASSR サーバにアクセスしたときに、またサーバとの対話全体で正確な話者認識を

促進し、ポータブルSSP装置とASSRサーバの間で正確な音声認識通信を行う機能を含む、ポータブルSSP装置が求められる。

【0007】本発明の実施例は、音声、沈黙及びバックグラウンド・ノイズの各信号を含むサウンドをアナログ信号に変換するマイク、アナログ信号をデジタル信号に変化するアナログ／デジタル・コンバータ、デジタル信号から音声を表すフィーチャ・ベクトル・データと、沈黙とバックグラウンド・ノイズの信号を表す特性データとを生成するDSP（デジタル信号プロセッサ）、遠隔地でのASSRサーバによる音声認識を目的に及び通信チャンネルを通してフィーチャ・ベクトル・データを表す信号を通信するために、音響またはデータの通信装置に接続するカプラを含むポータブルSSP装置を含む。カプラは、好適にはフィーチャ・ベクトル・データを音響信号に変換する音響カプラである。その場合、通信チャンネルも電話回線のようにアコースティック（acoustic）である。またカプラはデータ通信チャンネルを通して転送するためデジタル転送装置に接続ようにコネクタ、ポート、プロトコル等の対応するインタフェースを含む。

【0008】ポータブルSSP装置は、好適には、フィーチャ・ベクトル・データを暗号化する暗号化装置と、フィーチャ・ベクトル・データを圧縮するデータ圧縮装置を含む。ポータブルSSP装置は、好適には、ASSRサーバからの戻り信号を受信し処理する手段と、DSPによる処理を目的に戻り信号をデジタル戻りデータに変換する手段を含む。この好適な実施例ではDSPは更に、デジタル戻りデータを圧縮解除する手段とデジタル戻りデータを解読する手段を含む。

【0009】ポータブルSSP装置は更に、好適には話者、沈黙及びバックグラウンド・ノイズに関連する音響特性を含めて、通信チャンネルの伝達関数の予測を、好適には1組の予測基準信号を遠隔地でチャンネルに接続されたASSRサーバに送ることによって促進する手段を含む。ポータブルSSP装置は、そのデバイスに固有の暗号化キー・データまたは認証データを含めてデータを格納するメモリを含む。

【0010】本発明の他の好適な実施例は、マイクに向かって話されデジタル化された音声を処理し、音声を表すフィーチャ・ベクトル・データを生成するDSP（デジタル信号プロセッサ）と、電話回線またはデジタル・ネットワーク接続等の通信チャンネルに接続された通信装置に接続し、通信チャンネルを通して通信するためにフィーチャ・ベクトル・データを信号に変換するカプラと、ポータブルSSPから通信チャンネルを介して転送された信号を受信し、受信された信号を音声認識を目的に処理するため、通信チャンネルに接続された自動音声／話者認識（ASSR）サーバを持つポータブルSSP装置を持つASSRシステムを含む。

【0011】システムのASSRサーバは、登録または

認証のデータの格納モデルを含む。このモデルは加入者またはクライアントの登録時に作成される。ASSRサーバはまた音声認識のため、1組のボキャブラリと、言語モデル、隠れマルコフ・モデル（HMM、Hidden Markov Models）等他のモデルを格納する。ASSRサーバは、ポータブルSSP装置から受信された信号を処理し、処理された信号を格納モデルと比較する。

【0012】ここで好都合なことは、リモート話者認証機能があるとき、本発明の実施例に従ったシステムは、遠隔地のスマートカードまたは磁気カードの活動化／非活動化、またはパスワードもしくはPINコードの変更と再活動化の機能を提供することである。

【0013】更に、本発明の実施例に従ったリモート音声認識装置は、シグナリングの悪い、または歪みの大きい通信環境でも誤差が少なくボキャブラリの多い音声の認識が可能なASSR機能を提供する。

【0014】図1は、本発明の実施例に従ったポータブル音声信号前処理（SSP）装置99のブロック図を示す。マイク100は、ユーザが話した音声、沈黙及びバックグラウンド・ノイズを含むサウンドを受信し、サウンドをアナログ電気信号に変換するために用いられる。マイク100は、デジタル信号プロセッサ（DSP）120に転送して処理するためにマイク100からのアナログ電気信号をデジタル化信号に変換するアナログ／デジタル（A/D）・コンバータ110に接続される。DSP120は、好適にはプロセッサと、関連メモリ及びデータを処理しポータブルSSP装置99のデータの流れを制御する格納プログラムを含む。好適な実施例でDSP120の機能は、マイク100に向かって話された音声データをフィーチャ・ベクトルに前処理する機能、沈黙とバックグラウンド・ノイズのデータを処理して、通信チャンネルの伝達関数の確立または予測を補助する機能、及びポータブルSSP装置との間でデータの転送と受信を調整する、データを暗号化／解読する、必要に応じてデータを圧縮／圧縮解除する機能を含めた他のASSR機能を実行する機能を含む。これらの機能及び通信チャンネルの伝達関数の予測については以下で詳しく説明する。上に述べたDSPの機能または手法は当業者には知られており、一般に入手できる多数のDSPによりこれらの機能を実行できる。例えばTexas Instruments, Inc. のモデルTMS 32010、またはNippon Electric Co. のUPD 7720は、ここで述べる用途に適したDSPである。

【0015】DSP120によって処理されたデータは音響カプラ150に出力され、ここでデジタル・データはオーディオ信号に変換される。その場合、音響カプラからのオーディオ信号は、電話回線等のオーディオ通信チャンネルを通して転送するために、一般の電話機の受話器等のオーディオ通信装置で再生または話すことができる。音響カプラ150は本発明の好適な実施例に従

て、デジタル信号をオーディオ信号に変換する変換器と、一般の電話機の受話器のレシーバまたは送話口との接続に適したカブラを含む。例えば音響カブラの接続部分は、好適には、ポータブルSSP装置99を任意の一般電話機に取り付け、取り外すことができ、ユーザがどこにいても、電話から電話へポータブル装置の携帯性、可搬性を高めるように、一般電話機の送話口に確実に係合または接続できるゴム等の可撓物質で作られる。ポータブル装置は、好適にはバッテリーにより自己給電される。予想できるとおり、マイク、A/Dコンバータ、DSP、カブラ及びバッテリーを含むポータブル装置の基本コンポーネントは軽量コンパクトである。ポータブル装置はハンドヘルド型でもよく、或いはポケット電卓や財布のようにシャツのポケットに入れて携帯することもできる。

【0016】再び図1を参照する。暗号化装置130と圧縮装置140はポータブルSSP装置99のオプションのコンポーネントである。DSP120に接続された別々のチップもしくはモジュール、またはDSP130の内部プログラム等の暗号化装置130は、DSP120によって前処理されたデータを予め設定された暗号鍵で暗号化し、電話回線で信号の安全なやりとりを保証する。暗号鍵はユーザに依存するキー・コード等である。

【0017】好適にはDSP120は既知の信号をASSRサーバ200による処理のためにフィーチャ・ベクトルのストリームに追加する。この信号がASSRサーバ200によって知られており予想されている場合は、外部信号を簡単に抽出でき、サーバからクライアントへのプロセスの反転も簡単になる。例えば無線通信では、信号の分散スペクトル変調が用いられる。他のこれまでの暗号化法または暗号化アルゴリズムもこの暗号化／解読プロセスに使用できる。例えばBruce Schenierによる"Applied Cryptography", second edition, Wiley, 1996に述べられている暗号化アルゴリズムを参照されたい。

【0018】圧縮装置140はオプションであり、DSP120の外部に置かれたコンポーネントまたはチップでよく、或いはDSP120の格納プログラム機能として組み込むこともできる。圧縮装置140は前処理されたデータを、暗号化されているかされていないかと無関係に、電話回線を通して転送する前に圧縮し、電話回線の帯域幅が小さい場合に転送されるオーディオ信号により伝えられるデータの量を増やす。信号圧縮法は周知のとおりである。圧縮装置140は好適にはロスなく圧縮を行う。

【0019】マイク100からのデジタル化されたデータの前処理は、マイク100に向かって話された音声パターンを表すフィーチャ・ベクトルのシーケンスを生成するというこれまでの手法を含む。例えば米国特許番号第5544277号はこのような前処理の手法を1つ説

明している。またDSP120は沈黙及びバックグラウンド・ノイズのデータを処理して周囲データを作成し、このデータを電話回線を通して処理のためASSRサーバに転送する。ASSRサーバは、好適には中央に位置し、任意の電話機またはネットワークのステーションから簡単にアクセスできるように電話回線またはデジタル・ネットワークに接続される。

【0020】DSP120は、本発明の他の好適な実施例に従って、通信回線を通してASSRサーバから送られるオーディオ・データのポータブルSSP装置99による受信を調整する。ASSRサーバからのオーディオ・データは、ユーザに更に情報や指示を求めるプロンプトを含む。この実施例で音響カブラ150は、電話機の受話器の耳当てと係合するため音響カブラのトランスミッタと同じように設定され大きさが決められたレシーバを含む。受信されたメッセージを伝えるために、耳当てまたはカブラにはもう1つのスピーカを接続できる。DSP120は、オプションで、暗号化され圧縮されたデータがASSRサーバから送られた場合は、受信されたデータを圧縮解除し、また受信されたデータを解読できる格納プログラム機能を含む。音響カブラ150のレシーバ部は、DSP120で処理するために、受話器の耳当てからのオーディオ信号をデジタル信号に変換する。

【0021】図2に、本発明のポータブルSSP装置99を利用したシステムを示す。ASSRサーバ200は、好適にはポータブルSSP装置99から遠い中央に置かれ、電話回線を通して電話機からアクセスできる。接続が確立されると、ASSRサーバ200とポータブルSSP装置99、音声／話者認識装置との接続も確立される。ASSRサーバ200はポータブルSSP装置99と通信して、ユーザ側の沈黙及びバックグラウンド・ノイズを読取ることによって、通信チャネルの伝達関数を確立または予測する。通信チャネルの伝達関数の予測法は周知のとおりである。例えばB. C. Kuolによる"Automatic Control Systems", Prentice Hall, 1987及びJ. V. Candyによる"Signal Processing, the Modern Approach", McGraw Hill, 1988を参照されたい。本発明に従ったプロセス例を次に示す。DSP120は、転送または接続の初めに異なる信号を生成する。異なる信号は、既知の特性の上昇音 (ascending chirp) と下降音 (descending chirp)、等間隔の一連のサイン・トーン (sign tone)、及びホワイト・ノイズ信号である。信号間のオーダの期間と間隔は充分確立される。信号は入力トーンがサーバ200からポータブル装置99によって受信されるとすぐ生成される。これらの信号は2回生成される。一度はDSP120から、一度はマイク100からである。サーバ200によって受信される音響信号は既定の基準信号と比較され、チャネルとバックグラウンドの伝達関数が推測され、チャネルのシグネチャが作成される。このシグネチャは音声フィーチャ・ベクト

ルをデコンボルブ (deconvolve) するため用いられる。ASSRサーバ200は、好適には音声または話者を認識するためASSR格納プログラムを含む。例えば観測された音響特性をそれらに対応する状態にある場合に観測する可能性を記述し、最適経路での発音を認識することによって、lefeemesを表す隠れマルコフ・モデル (HMM) が用いられる。この可能性は、記述されたトレーニング・スピーチとHMMを整合させ、可能性を更新することによって調整することができる。例えば米国特許番号第5031217号及び第5276766号は、HMM及び他のモデルを用いた音声認識法について述べている。

【0022】本発明の実施例に従ったASSRシステムは、ユーザ側の沈黙及びバックグラウンド・ノイズの特性を示す。ASSRシステムは好適には、先に述べたように、バックグラウンド・ノイズを接続の初め、ユーザが沈黙している間、またはDSP120から基準信号が出される間にチェックする。バックグラウンド・ノイズはマイク100によって検出され、DSP120によって処理されてから基準信号の第2部分の後に送られる。ASSRサーバ200は、この情報からバックグラウンド・ノイズを推測し、受信済み入力ベクトルの処理に変更を加える (例えばケプストラム平均減法 (cepstral mean subtraction)、デコンボリューション等)。これに代えて、このような条件で調整された音声または話者の認識のための格納モデルも使用できる。同様に沈黙期間が検出され、単語間及びユーザが話す前の沈黙に関連付けられた音響特性が、好適にはバックグラウンド・ノイズを予測するために用いられる。

【0023】好都合なことは、本発明の実施例のASSRシステムが話者と音声の認識を実行することである。話者認識は、ASSRサーバにリモートにアクセスするために利用できる。好適には本発明のASSRシステム例は、通信されたデータをバックグラウンド・タスクとしてほぼ連続的に処理し、セッション全体で話者の存在を確認する。

【0024】本発明に従ったASSRの装置とシステムについて、通信装置を電話、通信チャネルを電話回線、ASSRサーバを電話回線に接続されたものとして示してきた。本発明では、通信チャネルはまた、インターネット、イントラネット及びローカル・エリア・ネットワーク等のデジタル通信チャネルでもあることを想定している。ASSRサーバ200は従って、電話によってかまたはデジタル・ネットワークによってアクセスすることができる。図3は、サーバ200とのインタフェースをとるためにマルチメディア・パーソナル・コンピュータ160をもつポータブル装置99の用途を示す。この実施例で受話器は、マイク及びスピーカを含むマルチメディア・コンピュータ160に代わっている。音響カプ

も近い位置に置かれ、音響カプ150のトランスミッタ部はコンピュータのマイクから最も近い位置に置かれる。コンピュータ160は、電話回線を通してデータを転送するためモデムを含むか、または内部サーバ300を通してサーバ200に接続できる。内部サーバ300はASSRサーバ200のようにASSR機能を含み、コンピュータ160と内部サーバ300の間にASSRシステムまたはASSR接続を確立できる。この実施例でマルチメディア・コンピュータはオーディオ通信装置として機能する。

【0025】これに代えてポータブルSSP装置99はコンピュータのI/Oポートとのコネクタを介してコンピュータに接続できる。この実施例で音響カプ150は使用せずに済ませることができる。DSP120からのデータはコンピュータのプロセッサに直接送ることができるからである。マルチメディア・コンピュータ160は、本発明の他の実施例に従って、上に述べたポータブルSSP装置99のコンポーネントを組み込んだポータブル・ラップトップ・マルチメディア・コンピュータ等である。よってポータブル・マルチメディア・ラップトップ・コンピュータは、ポータブルSSP装置99によって実行される全ての機能を実行することができる。ポータブル・ラップトップはモデムを通して、または内部サーバ300を通して、例えばTCP/IP、NetBEUI等のプロトコルによりASSRサーバ200に直接接続できる。

【0026】ポータブルSSP装置99の機能とコンポーネントは、本発明の他の実施例に従って、内蔵型セルラ・フォン (携帯電話) でもよい。従ってSSPセルラ・フォンはDSP120を含み、ASSRサーバ200との接続及びアクセスは、アナログまたはデジタル (例えばCDMA、GSM等) であるセルラ通信チャネルによる。

【0027】好都合なことは、ASSRサーバ200からのデータを受信、解読または圧縮解除するポータブル装置99の機能 (前述) により、解読または認証の機能を実行するために様々なデータを受信することができることである。例えば、現在及び次の通信の間に受信され、受信された信号が暗号化される暗号鍵、鍵の有効期間、鍵のPIN番号の有効期間等を確認するために受信されるクッキー、先に受信され、期限切れになるまでユーザを認証するために用いられるデジタル認証または認証、ノイズのマスキングまたはアンマスキングのため受信されるノイズ・パターン、スマートカードの組み込みチップでまたはクレジット・カード型カードの磁気バンドに、PINの有効期限の特性を示すクッキーと共にエンコードされるものの一部であるスマートカード用のPIN番号またはコード等である。受信されたデータはDSP200のメモリに格納できる。既知の手法によってDSP120の読取り/書込み/再プログラムを行うた

めにポータブル装置200に接続されるインタフェースを使用できる。

【0028】よって本発明のデバイス例と方法は、加入者またはクライアントのパスワード、ログオン、PIN、或いはまた暗号化／解読鍵の認証、リセット、または取り消しに使用できる。このようなタスクは、スマートカードまたは磁気バンドを持つカード上でリモートに実行できる。

【0029】スマートカード認証／PINリセット装置での本発明の実施例に従ったシステムの用途を図4に示す。

【0030】スマートカードの初期化：代表的なスマートカード・システムのスマートカードは、最初に機関の認証と秘密／公開鍵のセットで初期化されることによって、使用を目的としてアクティブにされる。登録サーバ410側では、通常は管理者は新しいスマートカードを管理者のスマートカードと共に挿入し、PIN番号を入力して新しいユーザの登録を承認する。管理者は次にユーザのスマートカードの初期化プログラムをアクティブにする。プログラムは通常、ユーザの秘密鍵及び公開鍵のセット、ユーザの名前、シリアル番号、スマートカード・シリアル番号等の認証用の情報をスマートカードにスタンプする。

【0031】図3のワークステーション220と同等でもよい登録サーバ410は、ユーザ・プロフィールを作成し、リクエスト秘密／公開鍵及び認証を生成し、情報をスマートカードにダウンロードする。登録サーバ410は次にユーザに識別のためにワークステーションのスピーカに向かって話すことを要求する。これらの音声メッセージは話者認識サーバ420に送られる。サーバ420は図1乃至図3のASSRサーバ200と同一である。音声メッセージ、認証及び一意のスマートカード・シリアル番号は、将来の認証及び他の用途のためにASSRサーバ200に関連付けられるデータベース、またはASSRサーバ200内のデータベースに保存される。

【0032】或いはまた認証が発行され、カスタム・プロフィールが作成されてから、登録サーバ410は登録データを保管のため保管サーバ(vault server)430及びディレクトリ・サーバ440にアップロードする。登録データはまたワークステーション450及びスマートカード・リーダ460にダウンロードすることもできる。ワークステーション450はポータブルSSP99のコンポーネントと機能を含むか、またはポータブルSSP99をワークステーション450に接続して話者認識サーバ420との話者／音声情報の通信を促進することができる。各機関に各スマートカードのために生成されるRSAキー・ペアがある。秘密鍵は、後でPINをリセットするときのためにカードに焼き込まれる。スマートカードは、将来使用するために内部(焼き込み)、

外部(カードへの印刷)の両方とも、一意のシリアル番号で初期化される。

【0033】上に述べたインストールまたは初期化の場合、ユーザがPINを忘れるか、PINが期限切れでリセットされていない場合は、PCまたはハンドヘルド・ポータブルSSP装置99とスマートカード・リーダ460(スマートカードの代わりに磁気バンド・カードが用いられる場合は磁気ライタ)によりASSRサーバ200とリンクしてPINリセットを要求することによって、次のようにして、PINをリモートにリセットまたは再設定することができる。

【0034】PCの場合、ユーザはASSRサーバ200との接続を(SSL V2等を通して)確立してスマートカードのPINの変更を要求する。ダイアログ・ボックスまたは音声プロンプトがユーザに提示され、ユーザのID、名前、スマートカード・シリアル番号等の入力求められる。ASSRサーバ200は、入力された情報をもとに格納された認証とユーザ・プロフィールにアクセスする。ASSRサーバ200は次にユーザに、認証のためプリセットされた音声メッセージでPCのスピーカに向かって話すことを求める。

【0035】アクセスされたユーザ・プロフィールとデータベースからの音声セグメントは、認証のためのユーザからの入力メッセージと比較される。ユーザには、確認プログラムへの入力を訂正するためにいくつかの機会を与えることができる。確認に問題がなく、ユーザが現在の有効なユーザならASSRサーバ200はスマートカード認証及び公開鍵を使用してPINリセット・コマンドを暗号化し、ユーザのPC及び関連するスマートカード・リーダに送る。ユーザのスマートカードはそこで自体の固有の秘密鍵を使用してRESET PINコマンドを解読する。

【0036】スマートカードの秘密鍵でRESET PINコマンドを問題なく暗号化できた場合、スマートカードはそこで活動化される。スマートカードには、スマートカードの焼き込まれた固有のシリアル番号をASSRサーバ200によって与えられた番号と比較するための回路を加えることができる。それらが同じなら、スマートカードはPIN RESET操作のため自体をアンロックする。

【0037】スマートカードが活動化されPINがリセットされると、ユーザは新しいPINをスマートカードに入力することができる。スマートカードは、任意の安全なトランザクションを目的にして利用できる。

【0038】電話によるPIN RESET: PCスピーカが利用できない場合でも、電話でASSRサーバ200を呼び出してスマートカードのPINリセットを行える。同じ話者認証手順により、サーバに、後に使用するために電子エンベロップで暗号化されたPIN RESETコマンドを作成することを指示することができる。

【0039】このPC及びスマートカード・リーダによ

り、ユーザは、ASSRサーバ200を通してスマートカード管理サーバとのSSL (Secure Socket Layer) セッションを確立し(ホスト認証だけで)、PINリセットを要求する。次にASSRサーバ200は、暗号化されたエンベロップをPCのスマートカード・リーダーに送る。スマートカードが電子エンベロップを受け取ると、スマートカードの秘密鍵により、エンベロップ内側のコマンドが解読される。正しい秘密鍵を持つスマートカードだけがPIN RESETコマンドを解読できる。スマートカードはまた、一意のスマートカード・シリアル番号をスマートカード管理サーバにより与えられた番号と比較する。データが一致すると、スマートカードは後のPIN INIT及びSETのためにリセットされる。

【0040】スマートカードのリモート無効化：スマートカードが失われるか盗まれた場合、ユーザはできるだけ早くスマートカードを無効にする必要がある。その場合、電話でASSRサーバ200を通してスマートカード管理サーバに接続し、スマートカードを同じ認証手順で無効化することができる。

【0041】上に示した、本発明の実施例に従ったシステム用途は、音声で確認するネットワークのアクセス・ログオンまたはパスワードにもあてはまる。

【0042】ネットワーク・ユーザは初期化プロセスを実行し、ASSRサーバ・データベースへの記録のためにプリセットされた1組の話者メッセージを記録する。ASSRサーバ200は、管理サーバとして、スマートカード・ユーザ認証について述べたように、話者認証プロセスを用い、ログオンID及びパスワードの全ての管理タスクを実行することができる。

【0043】スマートカードPINリセット・プロセスと同様に、ユーザがログオン・パスワードを忘れた場合またはパスワードが期限切れの場合、ユーザはASSRサーバ200を通じた認証により変更または再有効化を要求し、新しいパスワードまたは再有効化(reactivation)をユーザ側にダウンロードできる。他の用途には、電子マネー・カードまたはウォレット・カード(wallet card)との対話がある。ここではICチップに、カードに残っている金額に関する情報が含まれる。商品を購入するためにカードが使われたときには金額が少なくなる。本発明の実施例に従ったシステムは、上に述べたスマートカード再有効化プロセスと同様の手順によりマネー・カードをリフィル(refill)するために使用できる。

【0044】まとめとして、本発明の構成に関して以下の事項を開示する。

【0045】(1) 音声、沈黙、バックグラウンド・ノイズの各信号を含むサウンドをアナログ信号に変換するマイクと、アナログ信号をデジタル信号に変換するアナログ信号/デジタル・コンバータと、前記デジタル信号から前記音声を表すフィーチャ・ベクトル・データを生

成するデジタル信号プロセッサ(DSP)と、前記フィーチャ・ベクトル・データを音響信号に変換し、遠隔地で前記音声を認識するために、前記音響信号を通信チャネルを通して通信するために音響通信装置に接続する音響カプラと、を含む、ポータブル音声認識装置。

(2) 前記フィーチャ・ベクトル・データを暗号化する暗号化装置を含む、前記(1)記載の装置。

(3) 前記フィーチャ・ベクトル・データを圧縮するデータ圧縮装置を含む、前記(1)記載の装置。

(4) 前記音響カプラは、前記遠隔地からの前記通信チャネルを通じた戻り信号を受信し、前記戻り信号を前記デジタル信号プロセッサによる処理のためにデジタル戻りデータに変換する手段を含む、前記(1)記載の装置。

(5) 前記デジタル信号プロセッサは、前記デジタル戻りデータを圧縮解除する手段を含む、前記(4)記載の装置。

(6) 前記デジタル信号プロセッサは、前記デジタル戻りデータを解読する手段を含む、前記(4)記載の装置。

(7) 前記通信チャネルの伝達関数を予測する手段を含む、前記(1)記載の装置。

(8) 前記デジタル信号プロセッサは、前記遠隔地にて前記チャネルに接続されたサーバからデータを受信し処理する手段を含む、前記(1)記載の装置。

(9) 前記デジタル信号プロセッサは、前記装置に固有のキー・データを含むデータを格納するメモリを含む、前記(1)記載の装置。

(10) 前記装置を囲む沈黙とバックグラウンド・ノイズに関連した通信チャネル特性を把握する手段を含む、前記(1)記載の装置。

(11) デジタル信号プロセッサ(DSP)に向かって話されたデジタル化された音声を処理し、前記音声を表すフィーチャ・ベクトル・データを生成する前記デジタル信号プロセッサと、前記フィーチャ・ベクトル・データを表す信号を通信チャネルを通して通信するために通信装置に接続するカプラと、を含む、ポータブル音声信号プリプロセッサ(SSP)と、前記通信チャネルを介して前記ポータブル音声信号プリプロセッサから転送された前記フィーチャ・ベクトル・データを表す前記信号を受信し、前記音声または話者を認識するために前記信号を処理するために、前記通信チャネルとリモートに接続された自動音声/話者認識(ASSR)サーバと、を含む、音声認識システム。

(12) 前記自動音声/話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を処理し、処理された信号を認証データの前記格納モデルと比較することによって有効な加入者を認証する認証データ格納モデルを含む、前記(11)記載のシステム。

(13) 前記自動音声/話者認識サーバは、前記ポータ

ブル音声信号プリプロセッサから受信された前記信号を処理し、処理された信号をデータベースと比較することによって音声を確認するためのポキャブラリの前記データベースを含む、前記 ( 1 1 ) 記載のシステム。

( 1 4 ) 前記自動音声／話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を処理することによってユーザのパスワードを確認する、前記 ( 1 1 ) 記載のシステム。

( 1 5 ) 前記自動音声／話者認識サーバは、前記ポータブル音声信号プリプロセッサから受信された前記信号を 10 処理することによって呼び出し側を識別し、前記信号は、前記呼び出し側によって前記ポータブル音声信号プリプロセッサに話された音声から発する、前記 ( 1 1 ) 記載のシステム。

( 1 6 ) 前記ポータブル音声信号プリプロセッサは、前記自動音声／話者認識サーバから戻り信号を受信し処理する手段を含む、前記 ( 1 1 ) 記載のシステム。

( 1 7 ) 前記戻り信号は変更された認証データ及びプロンプトの 1 つを含む、前記 ( 1 6 ) 記載のシステム。

( 1 8 ) 前記ポータブル音声信号プリプロセッサは、前 20 記自動音声／話者認識サーバから受信されたデータを含むデータを格納する関連メモリを含む、前記 ( 1 1 ) 記載のシステム。

( 1 9 ) 前記通信チャネルを通した前記ポータブル音声信号プリプロセッサと前記自動音声／話者認識サーバの間の接続の伝達関数を予測する手段を含む、前記 ( 1 1 ) 記載のシステム。

( 2 0 ) 前記カブラは、前記フィーチャ・ベクトル・データを音響信号に変換する音響カブラであり、前記通信チャネルは音響通信チャネルである、前記 ( 1 1 ) 記載 30 のシステム。

( 2 1 ) 前記ポータブル音声信号プリプロセッサと前記自動音声／話者認識サーバは、通信セッション全体で話者認識を定期的に行う手段を含む、前記 ( 1 1 ) 記載のシステム。

( 2 2 ) クライアント／サーバ・システムの有効なクライアントを音声により認証する方法であって、登録データの複数のモデルを自動音声認識 ( A S S R ) サーバのメモリに格納するステップと、通信チャネルを通して前記自動音声認識サーバから遠く離れて配置されたポータ 40 ブル音声信号前処理 ( S S P ) 装置から前記自動音声認識サーバに接続するステップと、前記登録データを前記ポータブル音声信号前処理装置に話すステップと、前記ポータブル音声信号前処理装置の前記登録データを前処理してフィーチャ・ベクトルを作成するステップと、前記フィーチャ・ベクトルを表す信号を前記通信チャネルを通して前記自動音声認識サーバにより受信するステッ

プと、前記自動音声認識サーバから受信された前記信号を処理し、処理された信号を前記有効なクライアントを認証するために前記格納モデルと比較するステップと、を含む、方法。

( 2 3 ) 変更されたパスワード・データを前記自動音声認識サーバから前記ポータブル音声信号前処理装置に転送することによってユーザ・パスワードを変更するステップを含む、前記 ( 2 2 ) 記載の方法。

( 2 4 ) 変更されたスマートカード・データを前記自動音声認識サーバから前記ポータブル音声信号前処理装置に転送することによってスマートカードの P I N を変更するステップを含む、前記 ( 2 2 ) 記載の方法。

( 2 5 ) 前記変更されたスマートカード・データは前記自動音声認識サーバによって暗号化され、前記ポータブル音声信号前処理装置は前記スマートカードの P I N を変更するために前記変更されたスマートカード・データを解読する、前記 ( 2 4 ) 記載の方法。

( 2 6 ) 変更された P I N データを前記自動音声認識サーバから磁気カード・ライターと前記ポータブル音声信号前処理装置に転送することによって磁気カードの P I N を変更するステップを含む、前記 ( 2 2 ) 記載の方法。

【図面の簡単な説明】

【図 1】本発明の好適な実施例に従ったポータブル音声認識装置を示す図である。

【図 2】本発明の好適な実施例に従った図 1 の装置を持つシステムのブロック図である。

【図 3】本発明の好適な実施例に従った図 1 のポータブル装置を持つ他のシステムのブロック図である。

【図 4】本発明の好適な実施例に従ったスマートカード P I N 認証／リセット・システムの用法を示す図である。

【符号の説明】

9 9 ポータブル音声信号前処理 ( S S P ) 装置

1 0 0 マイク

1 1 0 アナログ・デジタル・コンバータ

1 2 0 デジタル信号プロセッサ ( P S P )

1 3 0 暗号化装置

1 4 0 圧縮装置

1 5 0 音響カブラ

1 6 0 マルチメディア・パーソナル・コンピュータ

2 0 0 サーバ

2 2 0 ワークステーション

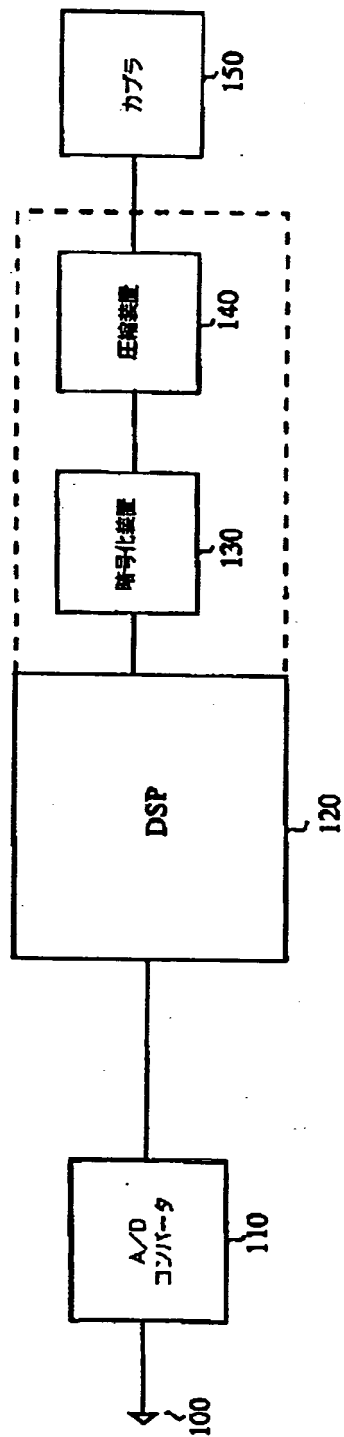
4 2 0 話者認識サーバ

4 3 0 保管サーバ

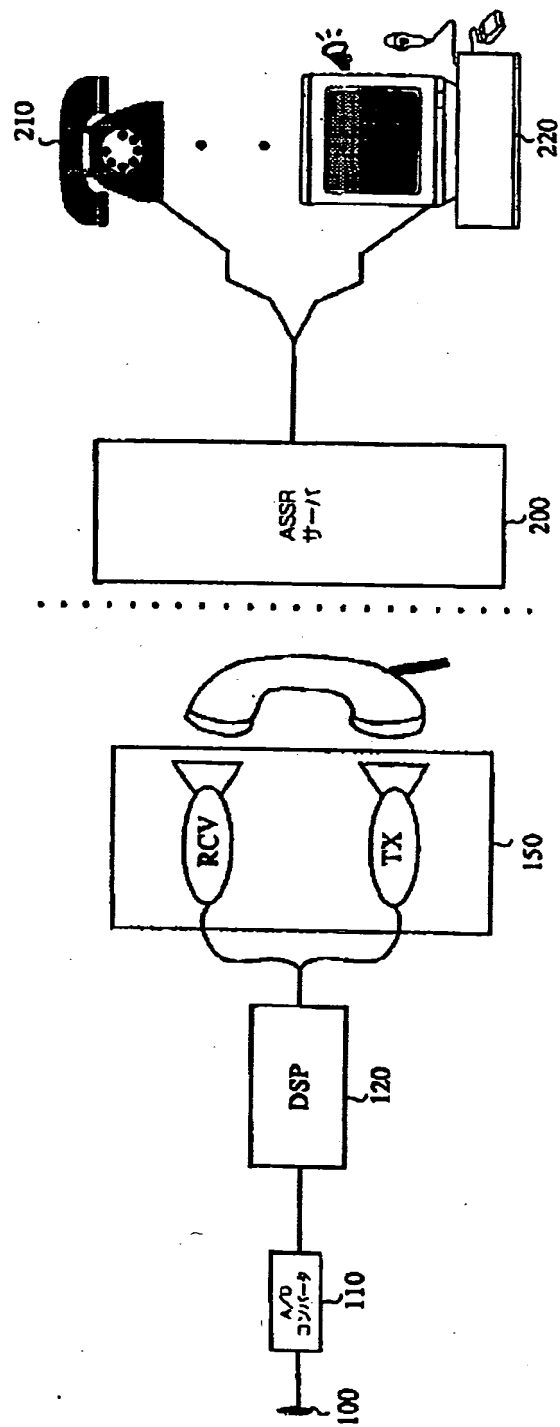
4 4 0 ディレクトリ・サーバ

4 6 0 スマートカード・リーダー

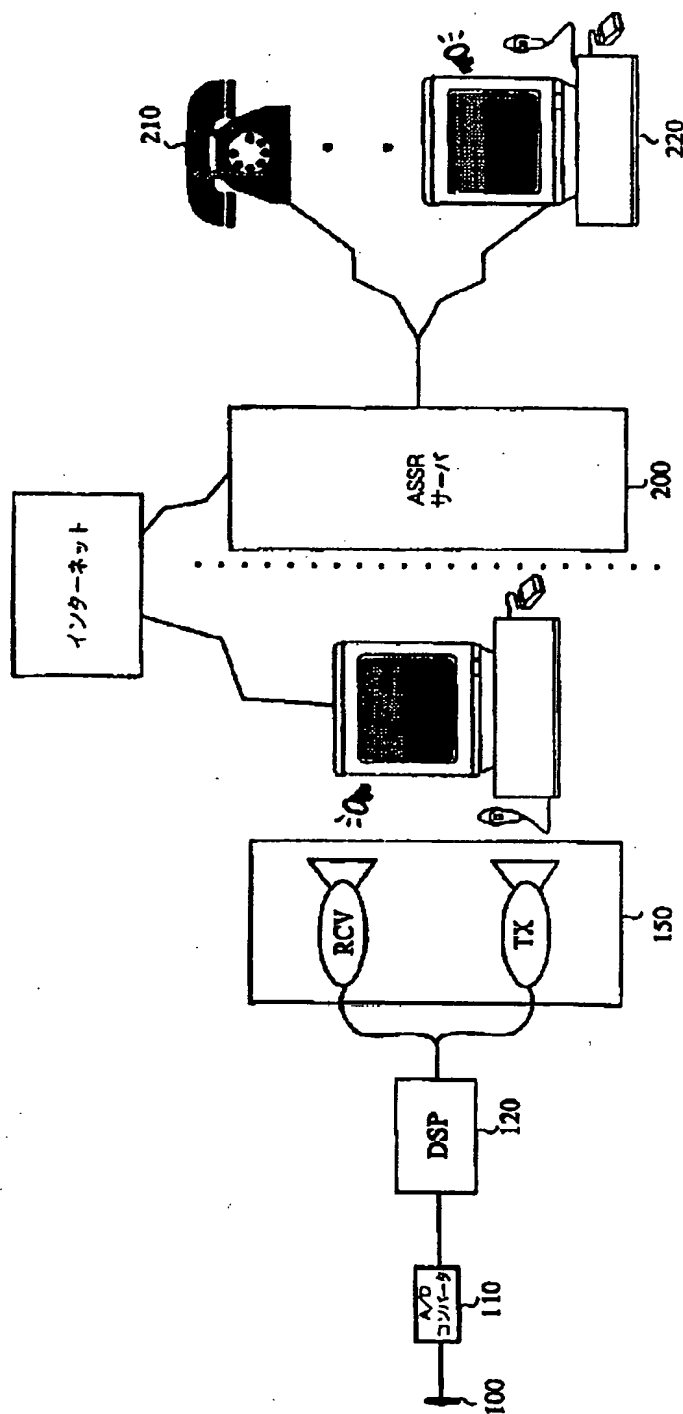
【図1】



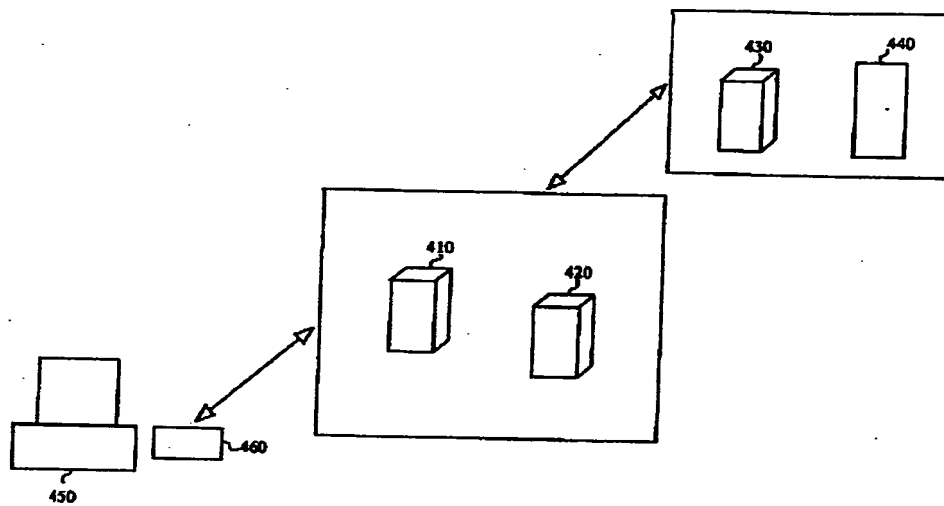
【図2】



【図 3】



【図 4】



フロントページの続き

(51) Int. Cl.<sup>4</sup>

識別記号

// H 0 4 M 3/42

F I

H 0 4 M 3/42

P

(72) 発明者 ステファン・ハーマン・メイス  
アメリカ合衆国06811、コネチカット州ダ  
ンバリー、シェルター・ロック・ロード  
157-25

(72) 発明者 ピーター・エス・ブーン  
アメリカ合衆国10589、ニューヨーク州ソ  
マーズ、サミット・サークル 10

(72) 発明者 カール・プロチロ  
アメリカ合衆国12498、ニューヨーク州ウ  
ッドストック、パーク・ドライブ 34